

Development of Sinhala Fonts in Sri Lanka

Gihan Dias

University of Moratuwa



How many letters are in the Sinhala alphabet?



- The Sinhala alphabet consists of 61 letters:
 - 18 vowels,
 - 41 consonants and
 - 2 semi-consonants

Source: SLS1134: Sri Lanka Standard - Sinhala character code for information interchange

Local-Language Computing in Sri Lanka



- Sri Lanka uses 3 languages
- Sinhala, Tamil, English
- Each uses a different script
- Most citizens know only either Sinhala or Tamil
- However, a fair amount of English literacy
- Most people know English alphabet
- Most people prefer to type in English
- Need to support Sinhala on Computers identified early
- Now all PCs and most phones support Sinhala and Tamil

Major Internet Sites support Sinhala and Tamil

The Sinhala Script



- The Sinhala language is a member of the Indic family of languages.
- Sinhala script is derived from Grantha and is similar to other south Indian Scripts
- An “abugida”, letter = consonant + vowel
- Letters are separated
 - Each syllable unit forms an individual visual unit or a glyph
- Two glyphs can tie up to form conjunct consonants (bendi akuru)
- Two **conjunct** letters – Yansaya and Rakaransaya



Sinhala Character Set

- Vowels

- අ ආ ඇ ඈ ඉ ඊ උ ඌ ඍ ඎ ඓ ඔ ඕ ඖ

- Semi-consonants

- ො and ෝ

- Consonants

ක බ ග ඝ ඩ ඞ

ච ඡ ජ කඨ කඳු ඳු ඒ

ට ඨ ඩ ධ ණ ඬ

ත ඵ ද ධ න ඳ

ප ඵ ඛ භ ම ඹ

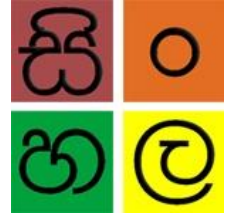
ය ර ල ව

Special features of the Sinhala Script



- Two additional vowels – අෑ අඳ
- Long e and o
- Sanngaka (half-nasal) letters – භ ඩ ද ඉ
- Theoretical letter – ඡ
- Unused letters – ඞ ඟ ජා
- Unofficial letter – ඳ

Development of Sinhala printing in the 20th Century



- Initially based on letterpress
- Movable lead type
- Many glyphs – each horizontal space is one piece of lead type
- Photo-typesetting – in 1970s
- Big, expensive machines

Development of Computer Typesetting in Sinhala



- Special-purpose word processors (e.g. Wang) in 1980s
- didn't become popular
- needed special displays
- Macintosh became popular for publishing
- as it had a graphical display
- First Sinhala fonts were developed for Mac
- by Wijeya Publications



Windows Fonts

- MS Windows became popular in the 1990s
- People started developing Sinhala fonts for Windows
- Basically displayed a particular Sinhala glyph for each key on keyboard
- Two types of keyboarding
 - “Phonetic” - each Sinhala letter is mapped to similar sounding English letter e.g., ක → k, ඛ → K
 - “Wijesekara” - similar to the well-known Sinhala typewriter keyboard
- In either case, the code assigned to each Sinhala letter is the same as that of the English letter at the same place on the keyboard



Windows Fonts (cont.)

- Though only two basic types of fonts, innumerable variations by different designers
- Slightly different codes
- May read incorrectly if wrong font is used
- Slightly different keyboarding
- Annoying
- Sometimes designers create special version of font to include a particular character

Windows Fonts



...	space	exclam	quotedbl	numbersig
															!	"	℥
dollar	percent	ampersand	quotesing	parenleft	parenright	asterisk	plus	comma	hyphen	period	slash	zero	one	two	three	four	five
₹	%	₹	'	()	X	+	,	-	.	/	0	1	2	3	4	5
six	seven	eight	nine	colon	semicolon	less	equal	greater	question	at	A	B	C	D	E	F	G
6	7	8	9	:	;	<	=	>	?	@	o	ඔ	ඒ	උ	ඌ	ඍ	ඎ
H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
ඹ	ය	ර	඼	ල	඿	඿	඿	඿	඿	඿	඿	඿	඿	඿	඿	඿	඿
Z	bracketlef	backslash	bracketrig	asciicircu	underscor	grave	a	b	c	d	e	f	g	h	i	j	k
ෆ	෇	'	෈	෉	්	෋	෌	෍	෎	ා	ෑ	ි	ී	ු	෕	ූ	෗
l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	braceleft	bar	braceright
ෘ	ෙ	ේ	ෛ	ෞ	෠	෡	෢	෣	෤	෥	෦	෧	෨	෩	෪	෵	෶
asciitilde	quotesing	florin	quotedbl	ellipsis	dagger	daggerdb	circumflex	perthousa	Scaron	guilsingll	OE	u008D	u008E	u008F
'																	
...	quoteleft	quoterigh	quotedbl	quotedbl	bullet	endash	emdash	tilde	trademark	scaron	guilsinglri	oe	u009D	u009E	Ydieresis	nbspace	exclamdo

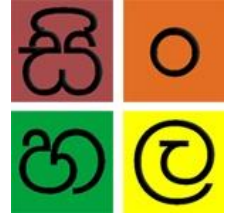
7-bit Phonetic Sinhala Font

7-bit Fonts



- 7-bit fonts assigned each key to a Sinhala letter or modifier
- As the Sinhala alphabet has more letters than English, some letters were missing
- Positioning could not be done properly, so documents did not look good

8-bit Fonts



- Assigned glyphs to all > 200 8-bit codes
- Needed to use special keyboard driver or use special alt-codes
- Full coverage and better glyphs

Features of Sinhala Script for Font Design



- Vowels and Sanngakas
- Similar-looking letters
- Letter modifiers
- Ascending and descending letters



Vowels and Sanngakas

- Sinhala has several letters which appear to have two or more components
- **But are not**
- Vowels: ආ ඈ ඉ ඊ උ ඌ ඍ ඎ ඏ ඐ එ ඒ ඓ ඔ ඕ
- Sanngakas: හ ඞ ඟ ච ජ
- Others: ක ඛ ඞ ජ



Similar-looking letters

- Some letters are distinguished only by minute differences
- e.g.
- න න
- හ හ
- බ බ
- ච ම
- සා සා
- A font needs to visually distinguish between them



Letter Modifiers

- Letters are modified by vowel signs which may occur not only after (post-modifier), but before the letter (pre-modifier)
- Upper- and Lower-modifiers change shape and position based on the length and shape of the base letter
- e.g: འ འྲ
- འི འོ འེ
- ཅ ཅྱ ཅཱ ཅུ
- ཇ ཇྱ

Ascending and Descending letters



- Some letters are tall, and adding an ispilla to them takes up more vertical space
- e.g. ఎ ఏ
- Some letters extend below the baseline and adding a papilla to them uses more space below
- e.g. ల ఱ

Unicode



- 7- and 8-bit fonts are glyph-based
- but
- Unicode stores text as a series of letters and modifiers
- Does not take into account
- letter shapes or
- modifier position

Unicode



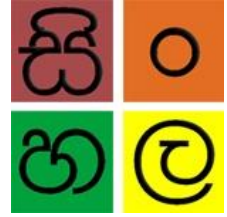
- In addition to glyphs, a Unicode font needs font rules
- and a font driver
- Difficult for traditional font designers
- Needs some programming knowledge

Sinhala in Unicode



- Each vowel is one Unicode character
- Each consonant is one character
- Each vowel sign and hal-lakuna is one character
- Always follows the consonant in code sequence
- Does not depend on the visual position of the sign
- Glyphs must be re-ordered to maintain correct visuals
- Special code sequences for Yansaya, Rakaransaya and Rephaya
- No special character for ඵ = ඵ + ඵ
- භ and භ have same vowel sign (and keyboarding)

Migration from Legacy fonts to Unicode



- Initially word processing, on-line newspapers, etc used legacy fonts
- Each document, website used whatever font it liked
- Websites included a link to download the appropriate font
- Now legacy-font WP documents are less common
- Pretty much *all* websites, including all newspapers, are in Unicode



වැඩිම ජාතන්

ලංකාදීප ONLINE

අන්තර්ජාලයට එකතු වූ ප්‍රථම සිංහල පුවත්පත | වෙබ්.ලංකාදීප.ලංකා

2011 නොවැම්බර් මස 02 වැනි බදද

- චසැණ පුවත්
- ශ්‍රීඩා විත්ති
- සයුරෙන් එතෙර
- අපේ උරුමය
- වීඩියෝ දර්ශන
- විශේෂාංග

රට සල්ලි අනුපාතය

චසැණ පුවත් >

ලැප්ටොප් බෑගයෙන් ඩොලර් මතුවේ

2011 නොවැම්බර් මස 02 18:40:17

(දීපා අධිකාරී)

අමෙරිකා ඩොලර් අනු දහසක් හෙවත් රුපියල් අනු ලක්ෂයකට අධික මුදලක් පාකිස්තානයේ කරවිව් නුවරට රැගෙන යාමට තැන් කළ පාකිස්තාන ජාතිකයෝ දෙදෙනෙක් අවසන් පර්යන්තයේදී අද සවස සැකපිට අත්අඩංගුවට ගනු ලැබූ..

0 - 784



Copyright@lankadeepa.lk

පැරිසිය ඡව් වර්ණයෙන් බැබළේ

2011 නොවැම්බර් මස 02 10:10:14

(පැරිසියේ සුනිල් ආනන්ද)

Microsoft (Monotype) Sinhala Font



- The first working font was introduced by Microsoft based on Monotype technology
- Called iskoola pota
- Very bad name
- Quite a nice font
- popular



Other Early Unicode Fonts

- Malithi Web – very basic font
- Unicode versions of legacy fonts
- Many used Iskoola Pota ruleset
- ICT Agency of Sri Lanka (ICTA) organised “Font Kitchen” by Muthu Nedumaran

Keyboard Methods

- Type-as-you-write
- Type-as-you-say
- Type as it is coded





Font Levels

- Sinhala fonts may be developed at 3 different levels.
- A level 1 font provides basic Sinhala support, and higher levels support more advanced features

Level 1 Fonts



- Intended to be used in mobile devices, etc.
- Represent commonly used vowels, consonants and consonants with modifiers
- Support Yansaya, Rakaransaya and Rephaya
- Support ఐతీ

Level 2 Fonts



- Usage: Level 2 fonts are intended for general applications such as documents, books, etc.
- Features: A level 2 font shall have all the features of level 1 font
- Support existing combination of Sinhala consonants with 'Rephaya'.

Level 3 Fonts



- Level 3 fonts are intended for advanced publications and printing applications,
- especially for supporting Pali and Sanskrit text and historical documents.
- Support Sinhala characters ඌ and ඌ
- Support all combinations of strokes with conjuncts including rephaya + ispilla combination.
- Support commonly used touching letters

ICTA Font Initiative



- ICTA set up an initiative to develop Sinhala (and Tamil) fonts
- Created Bhashitha and several display fonts
- <https://www.icta.lk/fonts/>

Issues in Developing Unicode Fonts



- Positioning vs glyph substitution
- Size of modifiers
- Glyph re-ordering

Current status



- Legacy fonts are disappearing
- Unicode has taken hold
- A few decent fonts
- Including Google's
- Not enough display fonts
- Less font design and development expertise
- Graphic design
- Font rule design

Way Forward



- Typoday will train new type designers
- Need to build on good ideas
- Should also have good technical knowledge

Thank You

Gihan Dias
gihan@uom.lk

